

AD-A172 007

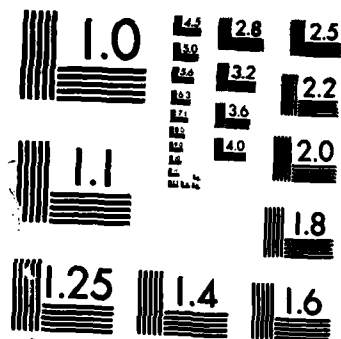
INFLUENTIAL NONNEGLECTIBLE PARAMETERS UNDER THE SEARCH
LINEAR MODEL(U) CALIFORNIA UNIV RIVERSIDE DEPT OF
STATISTICS S GHOSH 25 APR 86 TR-141 AFOSR-TR-86-0733
AFOSR-86-0048 F/G 12/1

1/1

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A172 007

AFOSR-TR- 86 - 0738

2

INFLUENTIAL NONNEGLECTIBLE PARAMETERS UNDER THE SEARCH

LINEAR MODEL*

BY

Subir Ghosh

Approved for public release;
distribution unlimited.

University of California, Riverside

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC

This technical report has been reviewed and is
approved for public release IAW AFR 190-12.

Distribution is unlimited.

MATTHEW J. KEEFER

Chief, Technical Information Division

Summary

In this paper some results useful in detecting the Influential
Nonnegligible parameters under the search linear model are presented.
An estimator of the number of nonnegligible parameters which are
significant and influential is also given.

AMS 1970 Subject Classifications: Primary and Secondary: 62J05,
62K05, 62K15, 62K99

Keywords and Phrases: Factorial Experiments, Influential
Parameters, Linear Models, Residuals, Search, Treatment.

DTIC
ELECTE
SEP 17 1986

B

*The work of the author is sponsored by the Air Force Office of
Scientific Research under Grant AFOSR-86-0048.

DTIC FILE COPY

1. Introduction

Consider a general factorial experiment with the design consisting of t treatments and corresponding to the u th treatment there are n_u (≥ 1) observations and $\sum_{u=1}^t n_u = N$. Let y_{uv} be the observation corresponding to the u th replication of the v th treatment and \bar{y}_u be the mean of all observations corresponding to the u th treatment. The model for this experiment is

$$\begin{aligned} E(\underline{y}) &= X_1 \underline{\beta}_1 + X_2 \underline{\beta}_2, \\ V(\underline{y}) &= \sigma^2 I, \\ \text{Rank } X_1 &= v_1, \end{aligned} \tag{1}$$

where $\underline{\beta}_1(v_1 \times 1)$ is a vector of specified lower order interactions and $\underline{\beta}_2(v_2 \times 1)$ is a vector of some or all of the higher order interactions, $X_1(N \times v_1)$ and $X_2(N \times v_2)$ are known matrices. It is known that K (very small compared to v_2) elements of $\underline{\beta}_2$ are nonzero and the other are zero; however the value of K and the nonzero elements of $\underline{\beta}_2$ are unknown. The problem is to search the nonzero elements of $\underline{\beta}_2$ and draw inferences on them in addition to the elements of $\underline{\beta}_1$. Such a model is called the search linear model and was introduced in Srivastava (1975). Suppose K_1 is an initial guess on K . Note the three possibilities $K_1 > K$, $K_1 = K$ and $K_1 < K$. We consider $\binom{v_2}{K_1}$ models

$$\begin{aligned} E(\underline{y}) &= X_1 \underline{\beta}_1 + X_2^{(1)} \underline{\beta}_2^{(1)}, \quad i=1, \dots, \binom{v_2}{K_1}, \\ V(\underline{y}) &= \sigma^2 I \\ \text{Rank}[X_1, X_2^{(1)}] &= v_1 + K_1, \end{aligned} \tag{2}$$

where $X_2^{(1)}(N \times K_1)$ is a submatrix of X_2 and $\beta_2^{(1)}(K_1 \times 1)$ is a subvector of β_2 . It can be seen from Srivastava (1975) that we in fact need $\text{Rank} [X_1, X_2^{(1)}, X_2^{(1')}] = (v_1 + 2K_1)$, for all $i \neq i'$. This implies that $N \geq (v_1 + 2K_1)$. In case $K_1 = K$, one of $\binom{v_2}{K_1}$ models is the correct model. If $K_1 > K$, then $\binom{v_2 - K}{K_1 - K}$ models out of $\binom{v_2}{K_1}$ models include the true model as a submodel in the expectation forms of the models. The methods discussed in this paper will not only identify K nonzero parameters but also find how many of them have significant effects and, finally, rank the significant nonnegligible parameters in the order of their influence on the fitted values. In case $K_1 < K$, the methods will identify from K_1 parameters the parameters which are significant and influential. We also propose an estimator of K in the Section 3.

In some industrial experiments, it is often easy to find replications ($n_u \geq 1$) in observations corresponding to a particular (the uth) treatments, see Taguchi and Wu (1985). There are also situations in industrial experiments where it is impossible to get replication in observations for a treatment, see Daniel (1976) and Box and Meyer (1985). The methods discussed in this paper consider both situations. In all Taguchi design methods, the higher order interactions (2-factor and higher order in most plans) are assumed to be zero. A few of those higher order interactions may be nonnegligible, significant and influential. The use of the search linear models may be a potential tool in improving upon the Taguchi design methods.



Codes	
Dist	Special
A-1	

2. Influential Nonnegligible Parameters.

Let $Z_i^{(1)}$ $((N-v_1-K_1) \times N)$ be such that $\text{Rank } Z_i^{(1)} = (N-v_1-K_1)$, $Z_i^{(1)} Z_i^{(1)'} = I$ and $Z_i^{(1)} [X_1, X_2^{(1)}] = 0$. Let $Z^{(1)} (K_1 \times N)$ be such that $\text{Rank } \begin{bmatrix} Z_i^{(1)} \\ Z^{(1)} \end{bmatrix} = (N-v_1)$, $Z^{(1)} Z^{(1)'} = I$, $Z^{(1)} Z_i^{(1)'} = 0$ and $Z^{(1)} X_1 = 0$. It can be seen that under the i th model in (2), the minimum variance unbiased estimator (MVUE) of $\beta_2^{(1)}$ is

$$\hat{\beta}_2^{(1)} = (Z^{(1)} X_2^{(1)})^{-1} Z^{(1)} \underline{y}. \quad (3)$$

In fact we can write $Z^{(1)'} = P_1 X_2^{(1)} D^{(1)}$, where $D^{(1)}$ is a nonsingular (and triangular) matrix so that $Z^{(1)} Z^{(1)'} = I$ and $P_1 = I - X_1 (X_1' X_1)^{-1} X_1'$. From the i th model in (2), the MVUE for β_1 is

$$\hat{\beta}_1^{(1)} = (X_1' X_1)^{-1} X_1' \underline{y} - (X_1' X_1)^{-1} X_1' X_2^{(1)} \hat{\beta}_2^{(1)}. \quad (4)$$

The fitted value of \underline{y} from the i th model in (2) is

$$\hat{\underline{y}}^{(1)} = X_1 \hat{\beta}_1^{(1)} + X_2^{(1)} \hat{\beta}_2^{(1)}. \quad (5)$$

The residuals from the i th model in (2) are

$$\begin{aligned} \underline{R}^{(1)} &= \underline{y} - \hat{\underline{y}}^{(1)} = P_1 (\underline{y} - X_2^{(1)} \hat{\beta}_2^{(1)}) \\ &= P_1 [I - X_2^{(1)} (X_2^{(1)'} P_1 X_2^{(1)})^{-1} X_2^{(1)'}] P_1 \underline{y}. \end{aligned} \quad (6)$$

The sum of squares due to error under the i th ($i = 1, \dots, \binom{v_2}{K_1}$) model in (2) is

$$SSE^{(1)} = \underline{R}^{(1)'} \underline{R}^{(1)} = \underline{y}' Z_i^{(1)'} Z_i^{(1)} \underline{y}. \quad (7)$$

The residuals under the model (1), when $\beta_2 = 0$, are

$$\underline{R}^{(0)} = \underline{y} - \hat{\underline{y}}^{(0)} = P_1 \underline{y}. \quad (8)$$

It can be seen that

$$P_1 = Z_1^{(1)'} Z_1^{(1)} + Z^{(1)'} Z^{(1)}. \quad (9)$$

Therefore, for $i = 1, \dots, \binom{v_2}{K_1}$,

$$SSE^{(0)} = \underline{R}^{(0)'} \underline{R}^{(0)} = SSE^{(1)} + \underline{y}' Z^{(1)'} Z^{(1)} \underline{y}. \quad (10)$$

For $i = 1, \dots, \binom{v_2}{K_1}$, we define

$$F^{(1)} = \frac{\underline{y}' Z^{(1)'} Z^{(1)} \underline{y} / K_1}{SSE^{(1)} / (N - v_1 - K_1)}. \quad (11)$$

Let $\hat{y}_u^{(1)}$ be the fitted value of the observation corresponding to the u th ($u = 1, \dots, w$) treatment under the i th model in (2). We write the sum of squares due to lack of fit as

$$SSLOF^{(1)} = \sum_{u=1}^w n_u (\bar{y}_u - \hat{y}_u^{(1)})^2, \quad (12)$$

and the sum of squares due to pure error as

$$SSPE = \sum_{u=1}^w \sum_{v=1}^{n_u} (y_{uv} - \bar{y}_u)^2. \quad (13)$$

For $i = 1, \dots, \binom{v_2}{K_1}$, we define

$$F_{LOF}^{(1)} = \frac{SSLOF^{(1)} / (w - v_1 - K_1)}{SSPE / (N - w)}. \quad (14)$$

Theorem 1. For $l \in \{1, \dots, \binom{V_2}{K_1}\}$, the following statements are equivalent.

- (a) $SSE^{(l)}$ is a minimum,
- (b) $F^{(l)}$ is a maximum,
- (c) $SSLOF^{(l)}$ is a minimum,
- (d) $F_{LOF}^{(l)}$ is a minimum,
- (e) The Euclidean distance between $\hat{y}^{(l)}$ and $\hat{y}^{(0)}$ is a maximum,
- (f) The square of the (sample) simple correlation coefficient between the elements of $\underline{R}^{(l)}$ and $\underline{R}^{(0)}$ is a minimum.

Proof. We have from (10) and (11) that

$$\left(\frac{K_1}{(N-v_1-K_1)} F^{(1)} \right) + 1 = \frac{SSE^{(0)}}{SSE^{(1)}}.$$

Noting that the numerator on the RHS of the above expression does not depend on i , we get the equivalence of (a) and (b). Again,

$$SSE^{(1)} = SSPE + SSLOF^{(1)},$$

and SSPE does not depend on i . Therefore (a) and (c) are equivalent.

From (14), the equivalence of (c) and (d) is clear. From (3), (6),

(8) and (9), it follow that

$$\begin{aligned} \underline{y}' \underline{Z}^{(1)'} \underline{Z}^{(1)} \underline{y} &= \hat{\underline{\beta}}_2^{(1)'} \underline{X}_2^{(1)'} \underline{Z}^{(1)'} \underline{Z}^{(1)} \underline{X}_2^{(1)} \hat{\underline{\beta}}_2^{(1)} \\ &= \hat{\underline{\beta}}_2^{(1)'} \underline{X}_2^{(1)'} \left[\underline{P}_1 - \underline{Z}_1^{(1)'} \underline{Z}_1^{(1)} \right] \underline{X}_2^{(1)} \hat{\underline{\beta}}_2^{(1)} \\ &= \hat{\underline{\beta}}_2^{(1)'} \underline{X}_2^{(1)'} \underline{P}_1 \underline{X}_2^{(1)} \hat{\underline{\beta}}_2^{(1)} \\ &= (\underline{R}^{(1)} - \underline{R}^{(0)})' (\underline{R}^{(1)} - \underline{R}^{(0)}) \\ &= (-\hat{\underline{y}}^{(1)} + \hat{\underline{y}}^{(0)})' (-\hat{\underline{y}}^{(1)} + \hat{\underline{y}}^{(0)}). \end{aligned} \tag{15}$$

The equivalence of (a) and (e) is now easy to see from (10) and (15).

It follows from (10) and (15) that $\underline{R}^{(1)'} \underline{R}^{(1)} = \underline{R}^{(1)'} \underline{R}^{(0)}$. We thus have

$$\frac{SSE^{(1)}}{\underline{R}^{(0)'} \underline{R}^{(0)}} = \frac{\underline{R}^{(1)'} \underline{R}^{(1)}}{\underline{R}^{(0)'} \underline{R}^{(0)}} = \frac{(\underline{R}^{(1)'} \underline{R}^{(0)})^2}{(\underline{R}^{(1)'} \underline{R}^{(1)}) (\underline{R}^{(0)'} \underline{R}^{(0)})} \quad (16)$$

= the square of the (sample) Simple correlation Coefficient between $\underline{R}^{(1)}$ and $\underline{R}^{(0)}$.

The equivalence of (a) and (f) is now clear from (16). This completes the proof of the theorem.

Proposition 1. Under the i th model in (2),

$$Z^{(1)} \underline{R}^{(1)} = \underline{0}. \quad (17)$$

Proof. It follows from (3) and (5) that

$$Z^{(1)} \underline{y} = Z^{(1)} X_2^{(1)} \underline{\beta}_2^{(1)} = Z^{(1)} \underline{y}.$$

This completes the proof.

We have

$$V(\underline{R}^{(1)}) = \sigma^2 P_1 \left[I - X_2^{(1)} (X_2^{(1)'} P_1 X_2^{(1)})^{-1} X_2^{(1)'} \right] P_1. \quad (18)$$

The residual in $\underline{R}^{(1)}$ are correlated and the question may be asked about the appropriateness in combining the elements of $\underline{R}^{(1)}$ in $SSE^{(1)}$.

If we take the transformed residuals as $Z_1^{(1)} \underline{R}^{(1)}$, we then have

$$E(Z_1^{(1)} \underline{R}^{(1)}) = \underline{0} \text{ and } V(Z_1^{(1)} \underline{R}^{(1)}) = \sigma^2 I. \quad (19)$$

The sum of squares of these transformed residuals is

$$\underline{R}^{(1)'} Z_1^{(1)'} Z_1^{(1)} \underline{R}^{(1)}.$$

Proposition 2. For $i = 1, \dots, \binom{v_2}{K_1}$,

$$SSE^{(i)} = \underline{R}^{(i)'} \underline{Z}_i^{(1)'} \underline{Z}_i^{(1)} \underline{R}^{(i)}. \quad (20)$$

Proof. We write the RHS using (9) as

$$\underline{R}^{(i)'} \underline{Z}_i^{(1)'} \underline{Z}_i^{(1)} \underline{R}^{(i)} = \underline{R}^{(i)'} \underline{P}_i \underline{R}^{(i)} - \underline{R}^{(i)'} \underline{Z}^{(1)'} \underline{Z}^{(1)} \underline{R}^{(i)}. \quad (21)$$

It can be checked that $\underline{P}_i \underline{R}^{(i)} = \underline{R}^{(i)}$. By using the Proposition 1, the rest of the proof is clear. This completes the proof.

Proposition 2 thus supports the use of $SSE^{(i)}$. Theorem 1 gives various interpretations of a search procedure, discussed in Srivastava (1975), of selecting $\underline{\beta}_2^{(i)}$ as the influential set of K_1 nonnegligible parameters.

We now denote

$$\begin{aligned} \underline{\beta}_2^{(i)'} &= [\beta_{21}^{(i)}, \dots, \beta_{2j}^{(i)}, \dots, \beta_{2K_1}^{(i)}], \\ \underline{x}_2^{(i)} &= [\underline{x}_{21}^{(i)}, \dots, \underline{x}_{2j}^{(i)}, \dots, \underline{x}_{2K_1}^{(i)}], \\ \underline{x}_2^{(ij)} &= \text{the matrix obtained from } \underline{x}_2^{(i)} \text{ by} \\ &\quad \text{deleting the } j\text{th column of } \underline{x}_{2j}^{(i)}. \\ \underline{x}_{12}^{(ij)} &= [\underline{x}_1, \underline{x}_2^{(ij)}], \\ \underline{P}_{12}^{(ij)} &= \underline{I} - \underline{x}_{12}^{(ij)} (\underline{x}_{12}^{(ij)'} \underline{x}_{12}^{(ij)})^{-1} \underline{x}_{12}^{(ij)'}, \\ \underline{z}_{1j}^{(i)} &= \frac{\underline{P}_{12}^{(ij)} \underline{x}_{2j}^{(i)}}{\sqrt{\underline{x}_{2j}^{(i)'} \underline{P}_{12}^{(ij)} \underline{x}_{2j}^{(i)}}}, \\ \underline{z}_0^{(i)'} &= [\underline{z}_{11}^{(i)}, \dots, \underline{z}_{1j}^{(i)}, \dots, \underline{z}_{1K_1}^{(i)}]. \end{aligned} \quad (22)$$

It can be seen that

$$\begin{aligned} \text{Rank} \begin{bmatrix} z_1^{(1)} \\ z_{1j}^{(1)'} \end{bmatrix} &= (N - v_1 - K_1 + 1), \\ z_{1j}^{(1)'} z_{1j}^{(1)} &= 1, \quad z_{1j}^{(1)'} z_1^{(1)} = 0, \\ \begin{bmatrix} z_1^{(1)} \\ z_{1j}^{(1)'} \end{bmatrix} X_{12}^{(1j)} &= 0, \quad \text{Rank } z_0^{(1)} = K_1. \end{aligned} \quad (23)$$

There exists a nonsingular (triangular) matrix $D_0^{(1)}$ such that

$$z^{(1)} = D_0^{(1)} z_0^{(1)}. \quad (24)$$

From (3) and (24), we have

$$\hat{\beta}_2^{(1)} = (z_0^{(1)} X_2^{(1)})^{-1} z_0^{(1)} y. \quad (25)$$

Now

$$z_0^{(1)} X_2^{(1)} = \text{diag} [z_{11}^{(1)'} X_{21}^{(1)}, \dots, z_{1j}^{(1)'} X_{2j}^{(1)}, \dots, z_{1K}^{(1)'} X_{2K}^{(1)}] \quad (26)$$

is a diagonal matrix. Thus

$$\hat{\beta}_{2j}^{(1)} = \frac{z_{1j}^{(1)'} y}{z_{1j}^{(1)'} X_{2j}^{(1)}}. \quad (27)$$

Let $\underline{R}^{(1j)}$, $i = 1, \dots, \binom{v_2}{K_1}$, $j = 1, \dots, K_1$, be the residuals obtained from i th model in (2) assuming $\beta_{2j}^{(1)} = 0$. Then the sum of squares due to error is

$$\text{SSE}^{(1j)} = \underline{R}^{(1j)'} \underline{R}^{(1j)} = (\underline{z}_{1j}^{(1)'} y)^2 + \text{SSE}^{(1)}. \quad (28)$$

We now define, for $i = 1, \dots, \binom{v_2}{K_1}$ and $j = 1, \dots, K_1$,

$$t^{(1j)} = \frac{\underline{z}_{1j}^{(1)'} \underline{y}}{\sqrt{\frac{SSE^{(1)}}{(N-v_1-K_1)}}} . \quad (29)$$

Proposition 3. For a fixed l in $\{1, \dots, (\frac{v_2}{K_1})\}$ and an m in $\{1, \dots, K_1\}$, the following statements are equivalent.

- (a) $SSE^{(lm)}$ is a minimum,
- (b) $t^{(lm)}$ is a maximum.

Proof. The proof can be easily seen from (28) and (29).

In the set $\underline{\beta}_2^{(l)}$ of influential nonnegligible parameters, $\beta_{2m}^{(l)}$ is the most influential nonnegligible parameters. The influential nonnegligible parameters may or may not have significant effects on observations.

3. Influential Significant Nonnegligible Parameters

We now assume the normality in (2) and therefore for $i = 1, \dots, \binom{v_2}{K_1}$, y independent $N(X_1\beta_1 + X_2^{(1)}\beta_2^{(1)}, \sigma^2 I)$. Under the null hypothesis $H_0: \beta_2^{(1)} = 0$, $F^{(1)}$ has the central F distribution with $(K_1, N - v_1 - K_1)$ d.f. and under the null hypothesis $H_0: \beta_{2j}^{(i)} = 0$, $t^{(ij)}$ has the central t distribution with $(N - v_1 - K_1)$ d.f. We now present a further development of a procedure suggested in Srivastava (1975).

Case I. If $\max_i F^{(i)} \leq F_{\alpha; K_1, N-v_1-K_1}$, we then conclude that there is no significant nonnegligible parameter. ($F_{\alpha; K_1, N-v_1-K_1}$ is the upper α percent point of the central F distribution with $(K_1, N-v_1-K_1)$ d.f.).

Case II. Suppose for $i = i_1, \dots, i_s$, we have $F^{(i)} > F_{\alpha; K_1, N-v_1-K_1}$. We denote for $j = 1, \dots, v_2$,

$$\partial_j = \text{the number of } i \text{ in } \{i_1, \dots, i_s\} \text{ for which } |t_{ij}| > \frac{t_{\alpha}}{2}, N-v_1-K_1.$$

Note that $0 \leq \partial_j \leq s$. We now arrange ∂_j 's in decreasing order of magnitude and write $\partial_{(1)} \geq \partial_{(2)} \geq \dots \geq \partial_{(v_2)}$. If there are at least K_1 nonzero $\partial_{(j)}$'s, we select the influential significant parameters as $\beta_{(1)}, \dots, \beta_{(K_1)}$, otherwise we pick the influential $\beta_{(j)}$'s corresponding to nonzero $\partial_{(j)}$'s (Note that the number of influential parameters is then less than K_1). The parameter $\beta_{(1)}$ is the most influential significant nonnegligible parameter. An estimator of the unknown K is \hat{K} = the number of nonzero ∂_j 's, $j = 1, \dots, v_2$.

4. Miscellaneous Results

4.a. Let us denote the unknown nonzero elements of $\underline{\beta}_2$ in (1) by $\underline{\beta}_{2c}(K \times 1)$ and the zero elements of $\underline{\beta}_2$ by $\underline{\beta}_{2d}((v_2 - K) \times 1)$, the corresponding columns in X_2 matrix are X_{2c} and X_{2d} . The unknown true expectation form of (1) is thus

$$E(\underline{y}) = X_1 \underline{\beta}_1 + X_{2c} \underline{\beta}_{2c}. \quad (30)$$

The expectation form of the i th model in (2) can be written as

$$E(\underline{y}) = X_1 \underline{\beta}_1 + X_{2c}^{(i)} \underline{\beta}_{2c}^{(i)} + X_{2d}^{(i)} \underline{\beta}_{2d}^{(i)}, \quad (31)$$

where $X_{2c}^{(i)}(N \times \gamma_1)$ is a submatrix of X_{2c} , $X_{2d}^{(i)}(N \times (K_1 - \gamma_1))$ is a submatrix of X_{2d} , $\underline{\beta}_{2c}^{(i)}(\gamma_1 \times 1)$ is a subvector of $\underline{\beta}_{2c}$ and $\underline{\beta}_{2d}^{(i)}((K_1 - \gamma_1) \times 1)$ is a subvector of $\underline{\beta}_{2d}$. Let $\underline{\beta}_{2c}^{*(i)}$ is the vector of elements in $\underline{\beta}_{2c}$ which are not in $\underline{\beta}_{2c}^{(i)}$ and $X_{2c}^{*(i)}$ is the matrix whose columns are in X_{2c} but not in $X_{2c}^{(i)}$. The following result, a counterpart of the result in (10) for the population, can be verified very easily.

Proposition 4. Under (30),

$$\begin{aligned} E(SSE^{(i)}) &= E(SSLOF^{(i)}) + \sigma^2(N-w) \\ &= \sigma^2(N-v_1-K_1) + \underline{\beta}_{2c}' X_{2c}' Z_1^{(i)'} Z_1^{(i)} X_{2c} \underline{\beta}_{2c} \\ &= \sigma^2(N-v_1-K_1) + \underline{\beta}_{2c}^{*(i)'} X_{2c}^{*(i)'} Z_1^{(i)'} Z_1^{(i)} X_{2c}^{*(i)} \underline{\beta}_{2c}^{*(i)} \quad (32) \\ &= E(SSE^{(0)}) - [\sigma^2 K_1 + \underline{\beta}_{2c}' X_{2c}' Z_1^{(i)'} Z_1^{(i)} X_{2c} \underline{\beta}_{2c}]. \end{aligned}$$

4.b. The model obtained from (2)

$$E(Z^{(1)} \underline{y}) = Z^{(1)} X_2^{(1)} \underline{\beta}_2^{(1)},$$

$$V(Z^{(1)} \underline{y}) = \sigma^2 I,$$

is called the pure search model (Srivastava (1976)). In fact, Srivastava (1976) considered a special form of $Z^{(1)}$.

4.c. The influential nonnegligible parameter may depend on noise, i.e., a parameter may be influential under one noise but may not be influential under another noise.

4.d. The replicated observations will surely improve the chances of detecting the correct influential nonnegligible parameters.

4.e. In presence of outliers in observations, one may combine residuals with unequal weights, or in other words, may use transformed residuals (see, Cook and Weisberg (1982)).

4.e.1. An example of transformed residual is the vector

$M^{(1)} \underline{R}^{(1)}$ where $M^{(1)} (N \times N)$ is a diagonal matrix whose u th diagonal element is $\left(1 / \sqrt{m_{uu}^{(1)}} \right)$ with $m_{uu}^{(1)}$ being the u th diagonal element of $\sigma^{-2} V(\underline{R}^{(1)})$.

4.e.2. Suppose the underlying design is robust against the unavailability of any single observation [see, Ghosh (1980)] in the sense that the estimation of $\underline{\beta}_1$ and $\underline{\beta}_2^{(1)}$ is possible under (2) when any single observation is unavailable. We find the predicted value of the u th observation from the remaining $(N-1)$ observations (i.e.,

by deleting the u th observation). The difference between the u th observation and its predicted value is called the u th predicted residual (using the idea of cross validation). It can be verified algebraically that the vector of predicted residuals is $[M^{(1)}]^2 \underline{R}^{(1)}$. The predicted residual sum of squares (PRESS) from the i th model under (2) is

$$PRESS^{(1)} = \underline{R}^{(1)'} [M^{(1)}]^4 \underline{R}^{(1)}$$

In presence of outliers, one may take $PRESS^{(1)}$ as an alternative to $SSE^{(1)}$.

BIBLIOGRAPHY

- Box, G. E. P. and Meyer, R. D. (1985) Studies in quality improvement I: dispersion effects from fractional designs. Tech. Report #2796. Mathematics Research Center, University of Wisconsin - Madison.
- Cook, R. D. and Weisberg, S (1982). Residuals and Influence in Regression. Chapman and Hall, New York,
- Daniel, C. (1976). Applications of Statistics to Industrial Experimentation. J. Wiley & Sons, New York.
- Ghosh, S. (1980). On robustness of optional balanced resolution V plans. J. of Statistical Planning and Inference, 4, 313-319.
- Ghosh, S. (1983). Influence observations in view of design and inference. Commun. Statist. - Theor. Meth., 12(14), 1675-1983.
- Srivastava, J. N. (1975). Designs for searching nonnegligible effects. In A Survey of Statistical Design and Linear Models (J. N. Srivastava, Ed.), pp. 507-519. North-Holland, Amsterdam.
- Srivastava, J. N. (1976). Some further theory of search linear models. In Contributions to Applied Statistics, Experimentia: Suppl. (W.J. Ziegler, Ed.), Vol. 22. Birkhäuser, Basel.
- Srivastava, J. N. and Mallenby, D. W. (1985). On a decision rule using dichotomies for identifying the nonnegligible parameter in certain linear models. J. of Multivariate Analysis, 16, 318-334.
- Taguchi, G. and Wu, Y. (1985). Introduction to Off-Line Quality Control. Central Japan Quality Control Association, Tokyo.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS A172 007	
2a. SECURITY CLASSIFICATION AUTHORITY NA		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for Public Release, Distribution Unlimited	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 141		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 86-0733	
6a. NAME OF PERFORMING ORGANIZATION University of California, Riverside	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM	
6c. ADDRESS (City, State and ZIP Code) DEPARTMENT OF STATISTICS UNIVERSITY OF CALIFORNIA RIVERSIDE, CA 92521		7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, DC 20332-6448	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-86-0048	
8c. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling AFB, DC		10. SOURCE OF FUNDING NOS.	
		PROGRAM ELEMENT NO. 6-1102F	PROJECT NO. 2304
11. TITLE (Include Security Classification) Influential Non-negligible Parameters Under The Search Linear Model		TASK NO. A5	WORK UNIT NO.
12. PERSONAL AUTHOR(S) SUBIR GHOSH			
13a. TYPE OF REPORT INTERIM	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day) 86-04-25	15. PAGE COUNT 15
16. SUPPLEMENTARY NOTATION SUBMITTED IN COMMUNICATIONS IN STATISTICS			
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB. GR.	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) In this paper some results useful in detecting the Influential Nonnegligible parameters under the search linear model are presented. An estimator of the number of nonnegligible parameters which are significant and influential is also given.			
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJOR BRIAN W. WOODRUFF		22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027	22c. OFFICE SYMBOL NM

END

DTIC

10-86